ISSN (Online) 2456-1290
International Journal of Engineering Research in Mechanical and Civil Engineering (IJERMCE)
Vol 1, Issue 6, October 2016

Assessment of Climate Change Impacts on Drought Pattern using Fuzzy C-Mean Clustering Approach – A Case Study of Rajasthan, India

[1]Ashutosh Sharma, [2] Manish Kumar Goyal
[1], [2] Department of Civil Engineering,
Indian Institute of Technology Guwahati
[1]ashutosh.sharma@iitg.ernet.in, [2] mkgoyal@iitg.ernet.in

Abstract: — **Drought is one of the most deadly natural phenomenons, which causes huge monetary losses around the globe. Changing climate and global warming is further anticipated to aggravate the drought scenarios. Assessment of climate change impacts on drought patterns plays important role in preparation of drought policies. Monthly rainfall data (from 1901 to 2000) is converted into standardized precipitation index (SPI) to quantify the drought intensity. Fuzzy c-mean (FCM) clustering approach is used to identify the homogeneous drought regions. Two cluster validity indices are used to validate the FCM parameters and obtain the optimal number of clusters. Finally, the whole data series is divided into four series to assess the impact of changing the climate on drought pattern. Drought pattern has shown a significant impact of climate change.**

Index Terms—**Drought regionalization, Fuzzy clustering, Standard precipitation index**

I. INTRODUCTION

In the field of hydrology, clustering algorithms are widely used for identification of the stations having similar hydro-climatological behavior [1], [2]. A cluster of homogeneous stations is generally called a region and hence, the process of identification of regions is called regionalization. In this study, homogeneous drought regions have been identified using a Fuzzy clustering approach. Fuzzy clustering offers some advantages over traditional hard clustering for hydrological applications [3]. Hard clustering assigns an object entirely to one cluster, whereas soft clustering allows the stations to be assigned to all clusters at the same time with a different degree of membership. This property of soft clustering provided additional information for the user to obtain more homogeneous clusters.

Droughts is an environmental condition of less precipitation over an extended period of time such as a season or a year.[4] It is quite different from other natural disasters in terms of its temporal scale and spatial extent. There are many indices available in the literature to quantify the intensity of the drought [5]. Standardized precipitation index (SPI), an index for quantification of meteorological drought intensity, is used in the present study [6]. SPI can be calculated for different scales such as 1 month, 3 months, 6 months, 12 month and even higher scales.

The main objective of this study is to assess the impact of changing climate on the drought regionalization. Climate change, along with global warming, has been affecting hydro-climatological processes across the globe [7], [8]. Many studies have reported the impact of climate change in the context of drought severity and pattern [9]. Rainfall pattern of last century has been analyzed for identification and assessment of climate change impacts on drought patterns of Rajasthan, India.

II. STUDY AREA AND DATA

a. Study area

Rajasthan, which is a western state of India, is taken as a study area for this study, considering the number of droughts that has occurred in this region in last century. The state is the largest state of India in terms of area, with an area of 342,000 km2 or around 10.4% of the total area of the country. Most parts of the state fall into the arid or semi-arid zone, which makes it highly vulnerable to droughts and water scarcity. The mean rainfall of the state is 574 mm but there is high spatial variation in amount of rainfall over the state. Rajasthan is classified as severely affected by droughts [10]. According to NIH, the frequency of deficient rainfall (75% of the normal or less) for the eastern Rajasthan is 1

ISSN (Online) 2456-1290
International Journal of Engineering Research in Mechanical and Civil Engineering (IJERMCE)
Vol 1, Issue 6, October 2016

in 3 years and for western Rajasthan is 1 in 2.5 years [10]. India has experienced 23 drought years between 1901 and 2002. Fig. 1 shows the location of study area (Rajasthan) and stations considered for this study.

b. Data

Monthly rainfall data at 32 stations spread across the study area was obtained from India water portal website (*http://www.indiawaterportal.org/met_data/*). Drought generally occurs over a long temporal scale, so data for 100 years, from January 1901 to December 2000, was used for this study. Latitude, longitude, altitude, maximum, minimum and average SPI values were used in cluster analysis. The numeric values of these variables vary too much because of the difference in the units, which may affect the clustering results. So, the data was normalized using the following transformation [1], [4]:

$$X_{i,j}^{N} = \frac{X_{i,j} - X_{i,\min}}{X_{i,\max} - X_{i,\min}} \quad \text{.....................} \ (1)$$

where $X_{i,j}^{N}$ denotes the normalized i^{th} attribute of j^{th} station; $X_{i,j}$ denotes i^{th} attribute of j^{th} station; $X_{i,\min}$ is the minimum of i^{th} attribute in all stations and $X_{i,\max}$ is the maximum of i^{th} attribute in all stations.

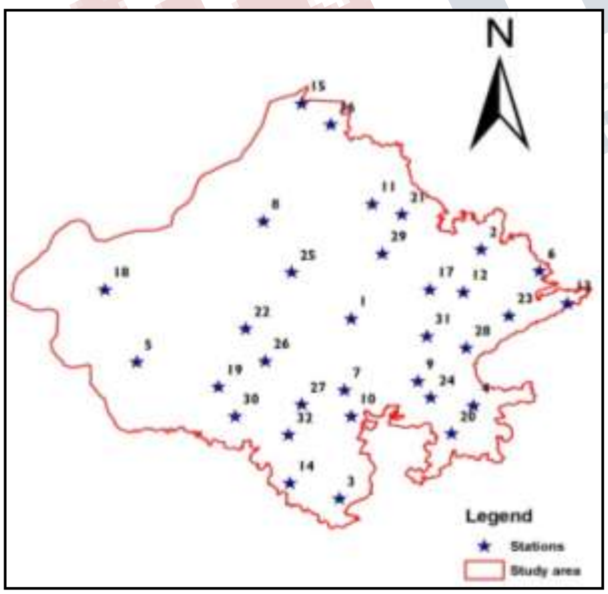

Fig. 1 Location of stations in the study area.

III. METHODOLOGY

Standardized Precipitation Index (SPI)

Standardized Precipitation Index (SPI) was developed by McKee [6], which can be defined on different time scales [11]. SPI is, generally, calculated for different time scales like 1, 3, 6, 12, 24 or 48 months. Classification of weather based on SPI values is given in Table 1.

Drought occurs when the value of SPI is continuously below zero and reaches an intensity of -1.0 or less. The drought ends when SPI value goes above zero. The duration of a drought is defined by its beginning and end, and intensity as the value of SPI. The magnitude of the drought is defined as the positive sum of the SPI values of all the months within the drought duration. Steps to compute SPI are explained in [4], [12].

Table 1: Weather classification based on SPI values

SPI Values	Weather classification
> 2	Extremely wet
1.5 to 1.99	Very wet
1.0 to 1.49	Moderately wet
-0.99 to 0.99	Near normal
-1.0 to -1.49	Moderately dry
-1.5 to -1.99	Severely dry
< -2	Extremely dry

Fuzzy c-mean (FCM) clustering

FCM algorithm was proposed by Dunn [13] and was further extended by Bezdek [14]. The algorithm iteratively optimizes a fuzzy objective function. Consider a cluster c having M objects in which Y_k is the data vector for the k^{th} object, $k = 1, 2, \ldots, M$. The fuzzy objective function is:

$$J(U,C) = \sum_{j=1}^{M} \sum_{i=1}^{c} u_{ik}^{\alpha} \| Y_k - C_i \|^2 \quad \text{................}(2)$$

where u_{ik} is degree of membership of k^{th} data point in i^{th} cluster, C_i is the center of i^{th} cluster, $\| Y_k - C_i \|^2$ is squared Euclidean distance of data vector k from the center of i^{th} cluster and α is called fuzzifier. Fuzzifier can have any value greater than 1, but generally, its value is set between 1 and 2.5.

ISSN (Online) 2456-1290
International Journal of Engineering Research in Mechanical and Civil Engineering (IJERMCE)
Vol 1, Issue 6, October 2016

Cluster validity

A number of cluster validity indices can be found in the literature, which are used to determine the optimal number of clusters (c) in the data set.[1], [4], [15], [12] For this study, we have used two cluster validity indices, Extended Xie-Beni index (V_{xb}) [16] and Kwon index (V_k) [15] for validating the number of clusters (c).

i. Extended Xie-Beni Index (V_{xb})

$$V_{XB,m}(U,V:X) = \frac{\sum_{i=1}^{c}\sum_{k}^{M}(u_{ik})^{\alpha} \| c_i - y_k \|^2}{M \min_{i \neq k} \| v_i - y_k \|^2} \quad \text{……….. (3)}$$

It is the ratio of compactness to the separation of the clusters. The minimum value of V_{xb} implies good clustering i.e. compact and well-separated clusters.

ii. Kwon Index (V_k)

The problem with V_{xb} is that its value monotonically decreases when number of clusters gets large. To solve this issue Kown [32] gave a new cluster validity index which contained an ad hoc punishing function in the numerator (2nd term in the numerator of the equation).

$$V_K(U,V:X) = \frac{\sum_{i=1}^{c}\sum_{k=1}^{M} u_{ik}^{\alpha} \| c_i - y_k \|^2 - \frac{1}{c}\sum_{i=1}^{c} \| c_i - \bar{c} \|^2}{\min_{i \neq k} \| c_i - y_k \|^2} \quad \text{4)}$$

IV. RESULTS AND DISCUSSION

The time series of rainfall data was divided into four series of 25 years namely Series 1 (1901-1925), Series 2 (1926-1950), Series 3 (1951-1975) and Series 4 (1976-2000). Monthly rainfall data was transformed into SPI of 6-month time scale for all four series. Maximum, minimum and mean SPIs for every series, along with three geographical attributes (latitude, longitude, and altitude) are used for clustering. Data for every series was normalized to eliminate the impact of attribute dimensions. To validate the number of clusters, two cluster validity indices, as described in previous section, are used.

Table 2: Selection of number of clusters based on V_k and V_{xb}.

No. of clusters	Series1 (S1)		Series2 (S2)		Series3 (S3)		Series4 (S4)	
	V_k	V_{xb}	V_k	V_{xb}	V_k	V_{xb}	V_k	V_{xb}
2	15.184	0.466	22.783	0.704	21.318	0.658	20.651	0.637
3	**12.348**	**0.372**	15.012	0.456	17.340	0.528	15.333	0.466
4	13.915	0.413	11.147	0.331	14.049	0.421	14.345	0.429
5	22.478	0.652	**8.102**	**0.234**	**13.710**	**0.404**	11.737	0.339
6	14.487	0.415	9.828	0.275	16.194	0.470	**8.228**	**0.227**

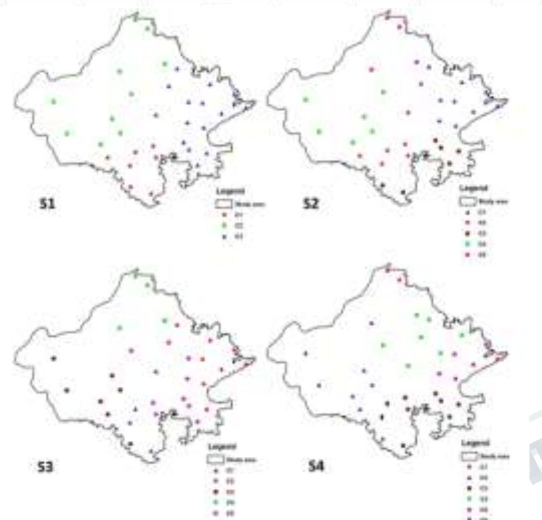

Fig. 2 Clusters for different series obtained using FCM. C1, C2, represents cluster numbers.

Table 2 shows the validity indices for different series. It can be seen that optimal number of cluster is increasing from S1 to S4, as per the minimum values of validity indices. The optimal number of clusters for S1 are 3 and for S2 and S3 are 5. A maximum number of clusters are found for later part of the century (S4), which could be due to the impact of climate change.

Fig. 2 shows the cluster obtained for different series. For S1, the state has three homogeneous drought regions. C2 represents the area which is severely affected by droughts. In S2 and S3, there are five homogeneous regions. The western part of the state is now further divided into two sub-regions. For S4, the state is divided into six regions, highest among the four series. This signifies that the drought pattern of the state is fluctuating. As in this study, we have only used the rainfall data, so these fluctuations in the drought pattern can be attributed to changes in the atmosphere/climate.

V. CONCLUSION

Climate change and global warming have affected the hydro-climatological processes. The impact of climate change on drought pattern in Rajasthan, India is examined using a Fuzzy c-means clustering approach. Number of clusters were validated using two indices, namely Extended Xie-Beni index, and Kwon index. Results have shown a significant impact of climate change on drought regionalization of the study area.

REFERENCES

[1] M. K. Goyal and V. Gupta, “Identification of homogeneous rainfall regimes in northeast region of India using fuzzy cluster analysis,” *Water Resour. Manag.*, vol. 28, no. 13, pp. 4491–4511, 2014.

[2] A. K. Mishra and V. R. Desai, “Drought forecasting using stochastic models,” *Stoch Env. Res Risk Assess 19 326–339*, pp. 326–339, 2005.

[3] A. R. Rao and V. V. Srinivas, “Regionalization of watersheds by fuzzy cluster analysis,” *J. Hydrol.*, vol. 318, no. 1–4, pp. 57–79, 2006.

[4] M. K. Goyal and A. Sharma, “A fuzzy c-means approach regionalization for analysis of meteorological drought homogeneous regions in western India,” *Nat. Hazards*, 2016.

[5] A. Zargar, R. Sadiq, B. Naser, and F. I. Khan, “A review of drought indices,” *Environ. Rev.*, vol. 19, no. NA, pp. 333–349, 2011.

[6] T. B. Mckee, N. J. Doesken, and J. Kleist, “The relationship of drought frequency and duration to time scales,” *AMS 8th Conf. Appl. Climatol.*, no. January, pp. 179–184, 1993.

[7] H. Li, C. X. S. Beldring, L. Merete, and T. Sharad, “Water Resources Under Climate Change in Himalayan Basins,” pp. 843–859, 2016.

[8] R. K. Pachauri and L. A. Meyer, *IPCC, 2014: Climate Change 2014: Synthesis Report. Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. 2014.

[9] S. V. R. K. Prabhakar and R. Shaw, “Climate change adaptation implications for drought risk mitigation: A perspective for India,” *Clim. Change*, vol. 88, no. 2, pp. 113–130, 2008.

[10] National Institutue of Hydrology (NIH), Report on “Draught Affected States.” Available on: *http://nihroorkee.gov.in/rbis/India_Information/draught.htm*

[11] A. M. Ticlavilca and M. McKee, “Multivariate Bayesian Regression Approach to Forecast Releases from a System of Multiple Reservoirs,” *Water Resour. Manag.*, vol. 25, no. 2, pp. 523–543, 2011.

[12] M. K. Goyal and A. Sharma, “Regionalization of Drought Prediction,” in *Handbook of Drought and Water Scarcity (HDWS)*, vol. 1, Francis and Taylor, CRC Group, 2016.

[13] J. C. Dunn, “A fuzzy relative of the ISODATA process and its use in detecting compact well-separated clusters,” *J. Cybern.*, vol. 3, no. 3, 1973.

[14] J. C. Bezdek, “Numerical taxonomy with fuzzy sets,” *J. Math. Biol.*, vol. 1, no. 1, pp. 57–71, 1974.

[15] M. K. Goyal and V. Gupta, “Impact of Climate Change on Regionalization Using Fuzzy Clustering,” in *Fourth International Conference on Soft Computing for Problem Solving, Advances in Intelligent Systems and Computing*, 2015, no. January, pp. 455–461.

[16] X. L. Xie and G. Beni, “A Validity Measure for Fuzzy Clustering,” *IEEE Trans Pattern Anal Mach Intell*, vol. 8, pp. 841–847, 1991.